

A QUANTITATIVE LOG-SOBOLEV INEQUALITY FOR A TWO PARAMETER FAMILY OF FUNCTIONS

E. INDREI AND D. MARCON

ABSTRACT. We prove a sharp, dimension-free stability result for the classical logarithmic Sobolev inequality for a two parameter family of functions. Roughly speaking, our family consists of a certain class of $\log C^{1,1}$ functions. Moreover, we show how to enlarge this space at the expense of the dimensionless constant and the sharp exponent. As an application we obtain new bounds on the entropy.

1. INTRODUCTION

1.1. Overview. Sobolev-type inequalities are central tools in analysis. The so-called logarithmic Sobolev inequalities appear in various branches of statistical mechanics, quantum field theory, and mathematical statistics. For example, the Gaussian log-Sobolev inequality is equivalent to Nelson's hypercontractive inequality and one may use log-Sobolev inequalities to show the stabilization of the Glauber-Langevin dynamic stochastic models for the Ising model with real spins, see for instance [13, 9].

Moreover, they are also useful in partial differential equations and Riemannian geometry. Indeed, they showed up in Perelman's work on the Ricci flow and the Poincaré conjecture [12]. While there is a large body of literature available on these inequalities, there are few corresponding stability results and this is currently an active area of research. Figalli, Maggi, and Pratelli [8] have recently addressed the stability problem for the anisotropic 1-log-Sobolev inequality; however, stability for the Gaussian log-Sobolev inequality – the classical version attributed to Stam [14], Federbush [5], and Gross [9] – is still open.

In this paper, we address this problem for a two parameter family of functions. Our approach involves techniques from optimal transport theory. Indeed, this theory has proven useful in producing sharp geometric and functional inequalities (see e.g. [7], [6]).

1.2. Main result. The classical Gaussian log-Sobolev inequality states that for smooth, positive functions

$$\int_{\mathbb{R}^n} f(x) \log f(x) d\gamma(x) - \|f\|_{L^1(d\gamma)} \log (\|f\|_{L^1(d\gamma)}) \leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f(x)|^2}{f(x)} d\gamma(x),$$

where $d\gamma := (2\pi)^{-n/2} e^{-|x|^2/2} dx$ is the standard Gaussian measure. The right hand side of the inequality is known as the *Fisher information* and is often denoted by $I(f)$ whereas the left hand side is the *entropy* and represented by $Ent(f)$. It is well-known that equality holds if and only if f is log linear (i.e. $f(x) = e^{a \cdot x + b}$). For $\epsilon > 0$ and $M > 0$, consider the family of functions:

$$\mathcal{F}(\epsilon, M) := \left\{ e^{-h} : (-1 + \epsilon) \leq D^2 h \leq M \right\},$$

and denote the *log-Sobolev deficit* by

$$\begin{aligned}\delta(f) &:= \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f(x)|^2}{f(x)} d\gamma(x) - \int_{\mathbb{R}^n} f(x) \log f(x) d\gamma(x) + \|f\|_{L^1(d\gamma)} \log(\|f\|_{L^1(d\gamma)}) \\ &= \frac{1}{2} I(f) - \text{Ent}(f).\end{aligned}$$

Note that $\delta \geq 0$ by the log-Sobolev inequality. The main result of this paper is the following theorem:

Theorem 1.1. *There exists an explicit dimensionless constant $C = C(\epsilon, M) > 0$ so that for all $f \in \mathcal{F}(\epsilon, M)$,*

$$W_2\left(f(\cdot)e^{-(\langle \mu, \cdot \rangle + |\mu|^2/2 + \log(m))} d\gamma, d\gamma\right) \leq C\delta(f/m)^{\frac{1}{2}},$$

where W_2 is the Wasserstein metric, $m = \|f\|_{L^1(d\gamma)}$, and μ is the barycenter of f .

Our theorem gives a quantitative way of measuring how far an admissible function is from attaining equality in the log-Sobolev inequality as measured with the Wasserstein metric. The proof of Theorem 1.1 is achieved by showing its equivalence to the following corollary (see §3):

Corollary 1.2. *There exists an explicit dimensionless constant $C = C(\epsilon, M) > 0$ so that for all $f \in \mathcal{F}(\epsilon, M)$ with unit mass and zero barycenter,*

$$W_2(f d\gamma, d\gamma) \leq C\delta(f)^{\frac{1}{2}}.$$

Although $\mathcal{F}(\epsilon, M)$ has a special structure, our results could be seen as a first step towards a sharp, general, dimension-free stability result for the Gaussian log-Sobolev inequality. In fact, by modifying our proof of Theorem 1.1, the class of admissible functions $\mathcal{F}(\epsilon, M)$ may be enlarged at the expense of the dimensionless constant and the sharp exponent. More specifically, thanks to a recent result of Kolesnikov [10], one may replace the upper L^∞ assumption on the Hessian of the logarithm of admissible functions with an L^r estimate.

Theorem 1.3. *If $\delta \leq 1$ and $r > 1$, then there exist explicit constants $C = C(\epsilon, M, n) > 0$ and $\beta = \beta(r) > 0$ so that for all*

$$f \in \tilde{\mathcal{F}}(\epsilon, M, r) := \left\{ e^{-h} : (-1 + \epsilon) \leq D^2 h, \int_{\mathbb{R}^n} \|(D^2 h + Id)_+\|^r f d\gamma \leq M \right\}$$

with unit mass and zero barycenter,

$$W_2(f d\gamma, d\gamma) \leq C\delta(f)^\beta.$$

Moreover, one may take $\beta = \frac{r-1}{2(2r-1)}$.

We remark that one may remove the unit mass and zero barycenter assumptions in Theorem 1.3 and prove an analogous result as in Theorem 1.1.

The paper is organized as follows: in §2, we collect some preliminary results from the literature which will be used in our proofs. In §3 we prove Theorems 1.1 and 1.3 and show that the $\frac{1}{2}$ exponent is sharp. Last, in §4 we show how to obtain bounds on the entropy in terms of the deficit (see Corollary 4.2) and derive an improved log-Sobolev inequality for our function class (see Remark 4.3).

2. PRELIMINARIES

Our proof of Theorem 1.1 exploits Cordero-Erausquin's optimal mass transfer proof of the log-Sobolev inequality [4]. For the reader's convenience and to simplify the presentation of our proof, we include his proof in this section along with statements of other results from the literature which will be useful for our purpose. We recall that given a smooth, positive function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ normalized to have unit mass with respect to the Gaussian measure $d\gamma$, Brenier's theorem yields the existence of a convex function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ such that its gradient $\nabla\phi$ is the optimal transport map between $fd\gamma$ and $d\gamma$: i.e.

$$\nabla\phi_{\#}(fd\gamma) = d\gamma$$

and

$$\int_{\mathbb{R}^n} |x - \nabla\phi(x)|^2 d\mu(x) = \inf_{T_{\#}(fd\gamma)=d\gamma} \int_{\mathbb{R}^n} |x - T(x)|^2 d\mu(x).$$

Moreover, ϕ satisfies $fd\gamma$ - a.e. the Monge-Ampère equation

$$f(x)e^{-|x|^2/2} = \det(D^2\phi)e^{-|\nabla\phi(x)|^2/2}.$$

For appropriate definitions from transport theory, we refer the reader to [15] (see also the introduction in [4] for a short and clear overview).

Theorem 2.1. (*log-Sobolev*) *Let f be a smooth, positive function on \mathbb{R}^n normalized to have unit mass with respect to the Gaussian measure $d\gamma$. Then,*

$$\int_{\mathbb{R}^n} f(x) \log f(x) d\gamma(x) - \|f\|_{L^1(d\gamma)} \log(\|f\|_{L^1(d\gamma)}) \leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f(x)|^2}{f(x)} d\gamma(x).$$

Proof as given by Cordero-Erausquin [4]. Without loss of generality, assume $\|f\|_{L^1(d\gamma)} = 1$. Let $\nabla\Phi$ be the Brenier map between $fd\gamma$ and γ , and set $\theta(x) := \Phi(x) - \frac{1}{2}|x|^2$ so that $\nabla\Phi(x) = x + \nabla\theta(x)$. It follows that $Id + D^2\theta \geq 0$, where Id is the identity matrix. The Monge-Ampère equation reads:

$$f(x)e^{-|x|^2/2} = \det(Id + D^2\theta)e^{-|x + \nabla\theta(x)|^2/2},$$

$fd\gamma$ - a.e. Taking the logarithm of both sides, the above equation may be rewritten as:

$$\begin{aligned} \log f(x) &= -\frac{1}{2}|x + \nabla\theta(x)|^2 + \frac{1}{2}|x|^2 + \log \det(Id + D^2\theta) \\ &= -x \cdot \nabla\theta(x) - \frac{1}{2}|\nabla\theta(x)|^2 + \log \det(Id + D^2\theta) \\ (2.1) \quad &\leq -x \cdot \nabla\theta(x) - \frac{1}{2}|\nabla\theta(x)|^2 + \Delta\theta(x), \end{aligned}$$

where the last inequality follows from the fact that $\log(1+t) \leq t$, for $t \geq -1$ (here, \log is the natural logarithm). Integrating with respect to $fd\gamma$ and using integration by parts, it follows

that

$$\begin{aligned}
\int_{\mathbb{R}^n} f \log f d\gamma &\leq \int_{\mathbb{R}^n} f[\Delta\theta - x \cdot \nabla\theta] d\gamma - \int_{\mathbb{R}^n} \frac{1}{2} |\nabla\theta|^2 f d\gamma \\
&= - \int_{\mathbb{R}^n} \nabla\theta \cdot \nabla f d\gamma - \int_{\mathbb{R}^n} \frac{1}{2} |\nabla\theta|^2 f d\gamma \\
&= - \int_{\mathbb{R}^n} \frac{1}{2} \left| \sqrt{f} \nabla\theta(x) + \frac{\nabla f(x)}{\sqrt{f}} \right|^2 d\gamma(x) + \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\gamma \\
&\leq \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} d\gamma.
\end{aligned}$$

□

Remark 2.2. Note that by the proof of Theorem 2.1 (more specifically, from (2.1)), if $\|f\|_{L^1(d\gamma)} = 1$, then

$$\delta(f) \geq \int_{\mathbb{R}^n} f \left(\Delta\theta - \log \det(Id + D^2\theta) \right) d\gamma.$$

Next, we state the following two theorems of Kolesnikov [10, Theorems 6.1 & 7.4]. The first generalizes Caffarelli's contraction theorem [3]:

Theorem 2.3. Let $\mu = e^{-V} dx$ and $\nu = e^{-W}$ be probability measures on \mathbb{R}^d and let $T = \nabla\Phi$ be the corresponding optimal transport map. If $D^2W \geq KId$, then for any $1 \leq r \leq \infty$,

$$K \|\Phi_{ee}^2\|_{L^r(\mu)} \leq \|(V_{ee})_+\|_{L^r(\mu)}.$$

Theorem 2.4. Let $\mu = e^{-V} dx$ and $\nu = e^{-W}$ be probability measures on \mathbb{R}^d and let $T = \nabla\Phi$ be the corresponding optimal transport map. If $D^2W \geq KId$, then for any $r \geq 1$,

$$K \left(\int_{\mathbb{R}^n} \|D^2\Phi\|^{2r} d\mu \right)^{\frac{1}{r}} \leq \left(\int_{\mathbb{R}^n} \|(D^2V)_+\|^r d\mu \right)^{\frac{1}{r}},$$

where $\|\cdot\|$ denotes the standard induced matrix norm.

Last, we shall make use of a well-known Poincaré type inequality for log convex measures, see [1, Theorem 2] (see also [2]):

Theorem 2.5. Let $\mu = e^{-V}$ be uniformly log convex with unit mass and $u \in H^1(\mu)$. Then,

$$\frac{1}{2} \left(\frac{p}{p-1} \right)^2 \left[\int_{\mathbb{R}^n} u^2 d\mu - \left(\int_{\mathbb{R}^n} |u|^{\frac{2}{p}} d\mu \right)^{2(p-1)} \cdot \left(\int_{\mathbb{R}^n} u^2 d\mu \right)^{\frac{2}{p}-1} \right] \leq \frac{2}{\lambda_1} \int_{\mathbb{R}^n} |\nabla u|^2 d\mu,$$

where $p \in (1, 2]$ and

$$\lambda_1 := \inf_{x \in \mathbb{R}^n, \xi \in \mathbb{S}^{n-1}} \langle D^2V(x)\xi, \xi \rangle > 0.$$

For our purposes, we will need the above theorem with $p = 2$.

Corollary 2.6. Let $\mu = e^{-V}$ be uniformly log convex with unit mass and $u \in H^1(\mu)$. Then

$$\int_{\mathbb{R}^n} |u - \bar{u}|^2 d\mu \leq C(\mu) \|\nabla u\|_{L^2(\mu)}^2,$$

where $\bar{u} := \int_{\mathbb{R}^n} u d\mu$ $C(\mu) := \frac{1}{\lambda_1}$.

3. PROOFS

3.1. L^∞ case.

Proof of Theorem 1.1. First, assume $\|f\|_{L^1(d\gamma)} = 1$ and that f has barycenter equal to zero with respect to $d\gamma$. Let $T = \nabla\Phi$ be the optimal transport between $fd\gamma$ and $d\gamma$ and set $\theta(x) := \Phi(x) - \frac{|x|^2}{2}$ (recall that Φ is convex). By Remark 2.2, we have

$$\delta(f) \geq \int_{\mathbb{R}^n} f \left(\Delta\theta - \log \det(Id + D^2\theta) \right) d\gamma.$$

We can express $\Delta\theta - \log \det(Id + D^2\theta)$ as

$$\begin{aligned} \Delta\theta - \log \det(Id + D^2\theta) &= \sum_{i=1}^n \lambda_i - \log \left(\prod_{i=1}^n (1 + \lambda_i) \right) \\ &= \sum_{i=1}^n (\lambda_i - \log(1 + \lambda_i)), \end{aligned}$$

where $\{\lambda_i\}_{i=1}^n$ are the eigenvalues of $D^2\theta$. Define $g : (-1, \infty) \rightarrow \mathbb{R}$ by $g(t) := t - \log(1 + t)$. For some $c > 0$ small enough, it follows that

$$g(t) \geq \phi(t) := c \min\{t^2, |t|\},$$

(this is easily seen by noting that g is quadratic at the origin and linear at infinity). Hence,

$$\begin{aligned} \delta(f) &\geq \sum_{i=1}^n \int_{\mathbb{R}^n} \phi(|\lambda_i|) f d\gamma \\ &\geq c \sum_{i=1}^n \left(\int_{\{x \in \mathbb{R}^n : |\lambda_i(x)| \geq 1\}} |\lambda_i| f d\gamma + \int_{\{x \in \mathbb{R}^n : |\lambda_i(x)| \leq 1\}} |\lambda_i|^2 f d\gamma \right). \end{aligned}$$

Let $\mu_i := 1 + \lambda_i \geq 0$ be the eigenvalues of $DT = D^2\Phi$ with corresponding eigenvectors v_i (recall that Φ is convex, so $\mu_i \geq 0$). It follows that

$$\mu_i(x) = \langle v_i(x), DT(x)v_i(x) \rangle \leq \sup_{e \in \mathbb{S}^{n-1}} \|\Phi_{ee}\|_\infty.$$

Now, we apply Theorem 2.3 with $V := h + \frac{|x|^2}{2}$ and $W := \frac{|x|^2}{2}$ and note that $D^2V = D^2h + I \leq (M+1)I$ to obtain

$$\sup_{e \in \mathbb{S}^{n-1}} \|\Phi_{ee}\|_\infty \leq \sup_{e \in \mathbb{S}^{n-1}} \sqrt{\|V_{ee}\|_\infty} \leq \sqrt{M+1}.$$

Therefore,

$$\|\lambda_i\|_\infty \leq C_M := \max\{1, \sqrt{M+1} - 1\}.$$

Set $E_i := \{x \in \mathbb{R}^n : |\lambda_i(x)| \geq 1\}$ and $\mu_f := fd\gamma$ so that

$$(3.1) \quad \int_{E_i} |\lambda_i|^2 d\mu_f \leq \|\lambda_i\|_\infty \int_{E_i} |\lambda_i| d\mu_f \leq C_M \int_{E_i} |\lambda_i| d\mu_f;$$

thus,

$$(3.2) \quad \delta(f) \geq (c/C_M) \int_{\mathbb{R}^n} \sum_{i=1}^n |\lambda_i|^2 d\mu_f = (c/C_M) \int_{\mathbb{R}^n} \|D(T(x) - x)\|_{HS}^2 d\mu_f,$$

where $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm. Now let $T = (T^1, T^2, \dots, T^n)$ and note that since $T_{\#}\mu_f = d\gamma$, we have

$$\int_{\mathbb{R}^n} (T(x) - x) d\mu_f = - \int_{\mathbb{R}^n} x d\mu_f(x) = 0.$$

By applying Poincaré (Corollary 2.6) to $T^i(x) - x_i$, we obtain

$$\begin{aligned} \int_{\mathbb{R}^n} |T(x) - x|^2 d\mu_f(x) &= \sum_i \int_{\mathbb{R}^n} |T^i(x) - x_i|^2 d\mu_f(x) \\ &\leq C(\mu_f) \sum_i \int_{\mathbb{R}^n} |\nabla(T^i(x) - x_i)|^2 d\mu_f(x) \\ &= C(\mu_f) \int_{\mathbb{R}^n} \sum_{i,j} |T_{x_j}^i - \delta_{ij}|^2 d\mu_f(x). \end{aligned}$$

Next, let $(a_{ij}(x))$ be the tensor $DT(x) - Id$ and note $a_{ij}(x) = T_{x_j}^i - \delta_{ij}$; in particular,

$$\int_{\mathbb{R}^n} \sum_{i,j} |T_{x_j}^i - \delta_{ij}|^2 d\mu_f(x) = \int_{\mathbb{R}^n} \|DT - Id\|_2^2 d\mu_f(x) = \int_{\mathbb{R}^n} \|DT - Id\|_{HS}^2 d\mu_f(x).$$

Combining this information with (3.2), we obtain

$$W_2^2(f d\gamma, d\gamma) = \int_{\mathbb{R}^n} |T(x) - x|^2 d\mu_f(x) \leq \frac{1}{c} C(\mu_f) C_M \delta(f).$$

Now $\lambda_1 \geq \epsilon$ since $f = e^{-h} \in \mathcal{F}(\epsilon, M)$ and so by applying Corollary 2.6, we obtain $C(\mu_f) \leq \frac{1}{\epsilon}$; setting $C = C(\epsilon, M) := \sqrt{\frac{1}{\epsilon}(\max\{1, \sqrt{1+M} - 1\})}$ completes the proof when $\|f\|_{L^1(d\gamma)} = 1$ and f has zero barycenter with respect to $d\gamma$. Next, assume

$$\mu = \int_{\mathbb{R}^n} x f(x) d\gamma \neq 0.$$

Define $\hat{f}(x) := f(x + \mu) e^{-\left(\mu \cdot x + \frac{|\mu|^2}{2}\right)}$. It is easy to see that

$$\int_{\mathbb{R}^n} x \hat{f}(x) d\gamma = \int_{\mathbb{R}^n} (x - \mu) f(x) d\gamma = 0,$$

and $\|f\|_{L^1(d\gamma)} = \|\hat{f}\|_{L^1(d\gamma)}$. Therefore, applying the previous argument to $\hat{f}/\|\hat{f}\|_{L^1(d\gamma)}$ yields

$$(3.3) \quad W_2((\hat{f}/\|\hat{f}\|_{L^1(d\gamma)}) d\gamma, d\gamma) \leq C \delta(\hat{f}/\|\hat{f}\|_{L^1(d\gamma)})^{\frac{1}{2}}.$$

Next we compute $\delta(\hat{f})$. To do this, suppose for the moment that $\|f\|_{L^1(d\gamma)} = 1$; then,

$$\begin{aligned}
\int_{\mathbb{R}^n} \hat{f} \log \hat{f} \, d\gamma &= \int_{\mathbb{R}^n} f(x + \mu) e^{-\left(\mu \cdot x + \frac{|\mu|^2}{2}\right)} \log f(x + \mu) \, d\gamma(x) \\
&\quad - \int_{\mathbb{R}^n} \left(\mu \cdot x + \frac{|\mu|^2}{2} \right) f(x + \mu) e^{-\left(\mu \cdot x + \frac{|\mu|^2}{2}\right)} \, d\gamma(x) \\
&= \text{Ent}(f) - \mu \cdot \int_{\mathbb{R}^n} x f(x + \mu) \, d\gamma(x + \mu) - \frac{|\mu|^2}{2} \\
&= \text{Ent}(f) - \frac{|\mu|^2}{2},
\end{aligned}$$

where we used that the barycenter of \hat{f} is zero with respect to $d\gamma$; moreover,

$$\begin{aligned}
\int_{\mathbb{R}^n} \frac{|\nabla \hat{f}|^2}{\hat{f}} \, d\gamma &= \int_{\mathbb{R}^n} \left| \frac{\nabla f(x + \mu) e^{-\left(\mu \cdot x + \frac{|\mu|^2}{2}\right)} - \mu f(x + \mu) e^{-\left(\mu \cdot x + \frac{|\mu|^2}{2}\right)}}{f(x + \mu) e^{-\left(\mu \cdot x + \frac{|\mu|^2}{2}\right)}} \right|^2 f(x + \mu) \, d\gamma(x + \mu) \\
&= \int_{\mathbb{R}^n} \left| \frac{\nabla f(x + \mu)}{f(x + \mu)} - \mu \right|^2 f(x + \mu) \, d\gamma(x + \mu) \\
&= \int_{\mathbb{R}^n} \left| \frac{\nabla f(x)}{f(x)} - \mu \right|^2 f(x) \, d\gamma(x) \\
&= I(f) - 2\mu \cdot \int_{\mathbb{R}^n} \nabla f(x) \, d\gamma + |\mu|^2 = I(f) - |\mu|^2,
\end{aligned}$$

where we used integration by parts in the last equality to deduce $\mu = \int_{\mathbb{R}^n} \nabla f(x) \, d\gamma$. The above considerations readily imply that if $\|f\|_{L^1(d\gamma)} = 1$, then $\delta(\hat{f}) = \delta(f)$, and the general case follows from the fact that δ is positively 1-homogeneous; combining these facts with (3.3) concludes the proof. \square

Remark 3.1. We note that the admissible functions are of the form e^{-h} , where h is semi-concave and semi-convex and the opening of the parabolas touching from above and below depend on the parameters ϵ and M . Therefore, the logarithm of the admissible functions have $C^{1,1}$ norms depending on these two parameters. In the proof, the upper bound on the Hessian of h was used to go from L^1 to L^2 (via Caffarelli/Kolesnikov), whereas the lower bound was used to apply Poincaré.

Remark 3.2. By considering a family of rescaled Gaussian measures it is not difficult to see that the exponent $\frac{1}{2}$ is sharp, see §3.3.

3.2. L^r case.

Proof of Theorem 1.3. Let $T = \nabla \Phi$ be the optimal transport between $f d\gamma$ and $d\gamma$ and set $\theta(x) := \Phi(x) - \frac{|x|^2}{2}$. By Remark 2.2, we have

$$\delta(f) \geq \int_{\mathbb{R}^n} f \left(\Delta \theta - \log \det(\text{Id} + D^2 \theta) \right) d\gamma.$$

We express $\Delta\theta - \log \det(I + D^2\theta)$ as

$$\begin{aligned}\Delta\theta - \log \det(I + D^2\theta) &= \sum_{i=1}^n \lambda_i - \log \left(\prod_{i=1}^n (1 + \lambda_i) \right) \\ &= \sum_{i=1}^n (\lambda_i - \log(1 + \lambda_i)),\end{aligned}$$

where $\{\lambda_i\}_{i=1}^n$ are the eigenvalues of $D^2\theta$. Define $g : (-1, \infty) \rightarrow \mathbb{R}$ by $g(t) := t - \log(1 + t)$ and

$$\phi(t) := \begin{cases} \frac{t^2}{2}, & -1 \leq t \leq 0 \\ t - \ln(1 + t), & t \geq 0. \end{cases}$$

Note that $\phi(t) = \phi(|t|)$ is convex and $g(t) \geq \phi(t)$. By Jensen's inequality, we obtain

$$\begin{aligned}\delta(f) &\geq \sum_{i=1}^n \int_{\mathbb{R}^n} \phi(|\lambda_i|) f \, d\gamma \\ &\geq \sum_{i=1}^n \phi \left(\int_{\mathbb{R}^n} |\lambda_i| f \, d\gamma \right).\end{aligned}$$

Since $\delta \leq 1$, it follows that

$$\phi \left(\int_{\mathbb{R}^n} |\lambda_i| f \, d\gamma \right) \geq c \left(\int_{\mathbb{R}^n} |\lambda_i| f \, d\gamma \right)^2,$$

for a small enough constant $c > 0$: in fact, $t - \log(1 + t) \geq ct^2$ for $0 \leq t \leq \frac{1-2c}{2c}$ and so if t^* satisfies $t^* - \log(1 + t^*) = 1$, we can pick c so that $t^* = \frac{1-2c}{2c}$ (i.e. $c = \frac{1}{2(1+t^*)}$); hence,

$$\delta(f) \geq c \sum_{i=1}^n \left(\int_{\mathbb{R}^n} |\lambda_i| f \, d\gamma \right)^2.$$

Now we apply Theorem 2.4 with $V := h + \frac{|x|^2}{2}$ and $W := \frac{|x|^2}{2}$ so that

$$\int_{\mathbb{R}^n} \|D^2\Phi\|^{2r} f \, d\gamma \leq \int_{\mathbb{R}^n} \|(D^2V)_+\|^r f \, d\gamma = \int_{\mathbb{R}^n} \|(D^2h + Id)_+\|^r f \, d\gamma \leq M.$$

Thus,

$$\begin{aligned}\int_{\mathbb{R}^n} |\lambda_i|^{2r} f \, d\gamma &\leq \int_{\mathbb{R}^n} \|D^2\Phi - Id\|^{2r} f \, d\gamma \\ &\leq \int_{\mathbb{R}^n} (\|D^2\Phi\| + 1)^{2r} f \, d\gamma \\ &\leq 2^{2r-1} \left(\int_{\mathbb{R}^n} \|D^2\Phi\|^{2r} f \, d\gamma + 1 \right) \\ &\leq 2^{2r-1} (M + 1) := C(r, M),\end{aligned}$$

Next, for $p > 2$, a standard interpolation inequality yields

$$\left(\int_{\mathbb{R}^n} |\lambda_i|^2 f \, d\gamma \right)^{\frac{1}{2}} \leq \|\lambda_i\|_{L^1(f \, d\gamma)}^\theta \|\lambda_i\|_{L^p(f \, d\gamma)}^{1-\theta},$$

where

$$\frac{1}{2} = \frac{(1-\theta)}{p} + \theta.$$

Thus, $p = \frac{1-\theta}{\frac{1}{2}-\theta}$ and we may pick $\theta = \frac{r-1}{2r-1}$ so that $p = 2r$. Hence,

$$\left(\int_{\mathbb{R}^n} |\lambda_i|^2 f d\gamma \right)^{\frac{1}{\theta}} \leq \left(\int_{\mathbb{R}^n} |\lambda_i| f d\gamma \right)^2 C(r, M)^{\frac{1-\theta}{r\theta}};$$

as θ depends on r , let $\tilde{C}(r, M) := \frac{1}{c} C(r, M)^{\frac{1-\theta}{r\theta}}$. Moreover, set $s := \frac{1}{\theta} > 2$ and $a_i := \int_{\mathbb{R}^n} |\lambda_i|^2 f d\gamma$ so that

$$\sum_i a_i^s \leq \tilde{C}(r, M) \delta(f).$$

Now by Hölder,

$$(3.4) \quad \left(\sum_i a_i \right)^s \leq \left(\sum_i a_i^s \right) n^{s-1} \leq \tilde{C}(r, M) n^{s-1} \delta(f).$$

Next,

$$(3.5) \quad \sum_i a_i = \int_{\mathbb{R}^n} \sum_i |\lambda_i|^2 f d\gamma = \int_{\mathbb{R}^n} \|D(T(x) - x)\|_{HS}^2 f d\gamma,$$

where $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm. As in the proof of Theorem 1.1, by applying Poincaré (Corollary 2.6) to $T^i(x) - x_i$,

$$\int_{\mathbb{R}^n} |T(x) - x|^2 d\mu_f(x) \leq C(\mu_f) \int_{\mathbb{R}^n} \|DT - Id\|_{HS}^2 d\mu_f(x).$$

Moreover, by combining this information with (3.5) and (3.4) we obtain

$$W_2^2(f d\gamma, d\gamma) = \int_{\mathbb{R}^n} |T(x) - x|^2 d\mu_f(x) \leq C(\mu_f) \left(C(r, M) n^{s-1} \right)^{\frac{1}{s}} \delta(f)^{\frac{1}{s}}.$$

Now $\lambda_1 \geq \epsilon$ since $f = e^{-h} \in \tilde{\mathcal{F}}(\epsilon, M)$ and so it follows from Corollary 2.6 that $C(\mu_f) \leq \frac{1}{\epsilon}$. This completes the proof and one may take $\beta := \frac{1}{2s} < \frac{1}{4}$. □

Remark 3.3. Note that the dimensional dependence came into play when we utilized Hölder's inequality in (3.4). Indeed, if $a_i \approx \text{constant}$, then

$$\frac{\sum_i a_i^s}{\left(\sum_i a_i \right)^s} \approx \frac{1}{n^{s-1}}.$$

Perhaps a different method may remove the dimension dependence; as we have seen in Theorem 1.1, this is possible under certain hypotheses (e.g. when one restricts the eigenvalues to be in L^∞).

3.3. Sharpness. In what follows, we show that the $\frac{1}{2}$ exponent in Theorem 1.1 is sharp by considering a family of rescaled Gaussians. First, we recall some basic facts: given $\mu \in \mathbb{R}^n$ and a symmetric, positive-definite matrix Σ , by setting

$$f(x) = \mathcal{N}(\mu, \Sigma) = \frac{1}{(2\pi)^{n/2}(\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right),$$

we have

$$\int_{\mathbb{R}^n} f(x) \, dx = 1$$

and

$$\int_{\mathbb{R}^n} x f(x) \, dx = \mu.$$

So if we define

$$f_a(x) := (2a + 1)^{n/2} e^{-a|x|^2},$$

since

$$\frac{1}{(2\pi)^{n/2}} f_a(x) e^{-\frac{|x|^2}{2}} = \left(\frac{2a + 1}{2\pi}\right)^{n/2} e^{-(a + \frac{1}{2})|x|^2} = \mathcal{N}\left(0, \frac{1}{2a + 1} I\right),$$

we readily obtain

$$\int_{\mathbb{R}^n} f_a(x) \, d\gamma(x) = 1$$

and

$$\int_{\mathbb{R}^n} x f_a(x) \, d\gamma(x) = 0.$$

In particular, given $\epsilon > 0$ and $M > 0$, we have that for all $a > 0$ small enough, $f_a \in \mathcal{F}(\epsilon, M)$. Moreover,

$$f_a \log f_a = \log(2a + 1)^{n/2} f_a - a|x|^2 f_a,$$

so

$$\int_{\mathbb{R}^n} f_a \log f_a \, d\gamma = \log(2a + 1)^{n/2} - a \int_{\mathbb{R}^n} |x|^2 f_a \, d\gamma,$$

and integrating by parts yields

$$\int_{\mathbb{R}^n} |x|^2 f_a \, d\gamma = n \int_{\mathbb{R}^n} f_a \, d\gamma + \int_{\mathbb{R}^n} x \cdot \nabla f_a \, d\gamma = n - 2a \int_{\mathbb{R}^n} |x|^2 f_a \, d\gamma,$$

which implies

$$\int_{\mathbb{R}^n} |x|^2 f_a \, d\gamma = \frac{n}{2a + 1}.$$

Thus, we may write the entropy as:

$$(3.6) \quad \text{Ent}(f_a) = \int_{\mathbb{R}^n} f_a \log f_a \, d\gamma = \log(2a + 1)^{n/2} - \frac{na}{2a + 1}.$$

Moreover, the Fisher information of f_a is given by:

$$(3.7) \quad \frac{1}{2} I(f_a) = \frac{1}{2} \int_{\mathbb{R}^n} \frac{|\nabla f_a|^2}{f_a} \, d\gamma = 2a^2 \int_{\mathbb{R}^n} |x|^2 f_a \, d\gamma = \frac{2na^2}{2a + 1}.$$

It is also not difficult to compute the Wasserstein distance between the two Gaussians $\mathcal{N}(0, \frac{1}{2a+1})$ and $\mathcal{N}(0, 1)$:

$$(3.8) \quad W_2\left(\mathcal{N}\left(0, \frac{1}{2a+1}I\right), \mathcal{N}(0, I)\right)^2 = n\left(\frac{1}{\sqrt{2a+1}} - 1\right)^2.$$

Therefore, by utilizing (3.6), (3.7), and (3.8) we deduce

$$\frac{\delta(f_a)^{1/2}}{W_2(f_a d\gamma, d\gamma)} = \frac{\sqrt{na - \frac{n}{2}\log(2a+1)}}{\sqrt{n}\left(1 - \frac{1}{\sqrt{2a+1}}\right)} = \frac{\sqrt{a - \frac{1}{2}\log(2a+1)}}{1 - \frac{1}{\sqrt{2a+1}}},$$

and repeated applications of l' Hôpital's rule yields that as $a \rightarrow 0$,

$$(3.9) \quad \frac{\delta(f_a)^{1/2}}{W_2(f_a d\gamma, d\gamma)} \rightarrow 1.$$

Since $\delta(f_a) \rightarrow 0$ as $a \rightarrow 0$, (3.9) implies that the exponent $\frac{1}{2}$ may not be replaced by something larger.

4. CONTROLLING THE ENTROPY

As an application of Corollary 1.2, we show how to obtain bounds on the entropy in terms of the deficit and barycenter. Let $d\gamma$ be the Gaussian measure and suppose that for a suitable class of functions f we have an estimate of the form:

$$(4.1) \quad W_2(f d\gamma, d\gamma) \leq C\delta^\alpha(f),$$

for some $\alpha \in (0, \frac{1}{2}]$. Thanks to Otto-Villani [11] (see also [4, Corollary 3]), we know that if f has unit mass with respect to the Gaussian,

$$(4.2) \quad Ent(f) \leq W_2(f d\gamma, d\gamma)\sqrt{I(f)} - \frac{1}{2}W_2^2(f d\gamma, d\gamma);$$

this inequality is known as the HWI inequality.

Lemma 4.1. *Suppose (4.1) holds for a suitable class of functions. Then*

$$Ent(f) \leq \tilde{C}(\delta^{\frac{1}{2}+\alpha}(f) + \delta^{2\alpha}(f)).$$

Proof. We simplify the notation in an obvious way. First, note that since all the quantities are non-negative

$$W\sqrt{I} = W\sqrt{2(\delta + E)} \leq \sqrt{2}W(\sqrt{\delta} + \sqrt{E}) = \sqrt{2}W\sqrt{\delta} + \sqrt{2}W\sqrt{E},$$

(using $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$). Now,

$$W\sqrt{E} = \frac{W}{\epsilon}(\epsilon\sqrt{E}) \leq \frac{W^2}{2\epsilon^2} + \frac{\epsilon^2}{2}E,$$

(using $ab \leq \frac{1}{2}(a^2 + b^2)$) and thus,

$$W\sqrt{I} \leq \sqrt{2}W\sqrt{\delta} + \sqrt{2}\left(\frac{W^2}{2\epsilon^2} + \frac{\epsilon^2}{2}E\right).$$

Hence, an application of (4.2) yields

$$E \leq \sqrt{2}W\sqrt{\delta} + \sqrt{2}\left(\frac{W^2}{2\epsilon^2} + \frac{\epsilon^2}{2}E\right) - \frac{1}{2}W^2 = \sqrt{2}W\sqrt{\delta} + \left(\frac{\sqrt{2}}{2\epsilon^2} - \frac{1}{2}\right)W^2 + \frac{\sqrt{2}\epsilon^2}{2}E;$$

using (4.1) we have

$$(1 - c\epsilon^2)E \leq C(\delta^{\frac{1}{2}+\alpha} + \delta^{2\alpha}).$$

Picking ϵ sufficiently small completes the proof. \square

Corollary 4.2. *There exists an explicit dimensionless constant $\bar{C} = \bar{C}(\epsilon, M) > 0$ so that for all $f \in \mathcal{F}(\epsilon, M)$,*

$$Ent(f) \leq \bar{C}\delta(f) + \frac{1}{2}|\mu(f)|^2,$$

where $\mu(f)$ is the barycenter of f with respect to $d\gamma$.

Proof. Given $f \in \mathcal{F}(\epsilon, M)$, let

$$\hat{f}(x) := f(x)e^{-\left(\langle \mu(f), x \rangle + |\mu(f)|^2/2 + \log(\|f\|_{L^1(d\gamma)})\right)},$$

where $\mu(f)$ is the barycenter of f with respect to $d\gamma$. Note that by Theorem 1.1,

$$W_2(\hat{f}d\gamma, d\gamma) \leq C\delta(\hat{f}/\|f\|_{L^1(d\gamma)})^{\frac{1}{2}}.$$

Hence, Lemma 4.1 implies

$$Ent(\hat{f}/\|f\|_{L^1(d\gamma)}) \leq \bar{C}\delta(\hat{f}/\|f\|_{L^1(d\gamma)}),$$

and so

$$Ent(\hat{f}) \leq \bar{C}\delta(\hat{f}).$$

From the proof of Theorem 1.1, we know $\delta(\hat{f}) = \delta(f)$ and $Ent(\hat{f}) = Ent(f) - \frac{1}{2}|\mu(f)|^2$, and this yields the result. \square

Remark 4.3. *Let $f \in \mathcal{F}(\epsilon, M)$ and consider \hat{f} as in the proof of Corollary 4.2. Since \hat{f} has zero barycenter,*

$$Ent(\hat{f}) \leq \bar{C}\delta(\hat{f}).$$

Thus,

$$Ent(f) \leq \frac{\bar{C}}{2(\bar{C}+1)}I(f);$$

now, as $C := \frac{\bar{C}}{2(\bar{C}+1)} < \frac{1}{2}$, this improves the constant in the log-Sobolev inequality for functions in $\mathcal{F}(\epsilon, M)$ with unit mass and zero barycenter with respect to $d\gamma$. More generally, since $Ent(\hat{f}) = Ent(f) - \frac{1}{2}|\mu(f)|^2$ and $I(\hat{f}) = I(f) - |\mu(f)|^2$ (see e.g. the proof of Theorem 1.1), it follows that

$$(4.3) \quad Ent(f) \leq CI(f) + ((1/2) - C)|\mu(f)|^2.$$

We note that as $M \rightarrow \infty$ or $\epsilon \rightarrow 0$, $\bar{C} = \bar{C}(\epsilon, M) \rightarrow \infty$, and so when we enlarge our function space in this way, $C \rightarrow \frac{1}{2}$ – the sharp log-Sobolev constant for general functions. Indeed, (4.3) measures the improvement in the log-Sobolev constant for the class $\mathcal{F}(\epsilon, M)$ in terms of the barycenter.

Acknowledgments. We wish to thank Alessio Figalli for suggesting the problem and for lively discussions on this topic. E. Indrei was supported by a departmental fellowship for graduate studies at the University of Texas at Austin. D. Marcon was supported by the UT Austin-Portugal partnership through the FCT doctoral fellowship SFRH/BD/33919/2009. Most of this work was completed while the authors were participating in the program “Concentration month on nonlinear elliptic PDEs” at the University of Chicago – the excellent research environment is kindly acknowledged.

REFERENCES

- [1] A. Arnold and J. Dolbeault, Refined Convex Sobolev Inequalities, *J. Funct. Anal.*, Volume 225, Issue 2, 15 August 2005, Pages 337 - 351.
- [2] J.P. Bartier and J. Dolbeault, Convex Sobolev inequalities and spectral gap, *C. R. Acad. Sci. Paris*, Ser. I 342 (2006) 307-312.
- [3] L.A. Caffarelli, Monotonicity properties of optimal transportation and the FKG and related inequalities, *Comm. Math. Phys.*, 214 (3), (2000), 547-563.
- [4] D. Cordero-Erausquin, Some applications of mass transport to Gaussian type inequalities, *Arch. Rational Mech. Anal.* 161 (2002), 257-269.
- [5] P. Federbush, A partially alternate derivation of a result of Nelson, *J. Physics* 10 (1969), 50-52.
- [6] A. Figalli and E. Indrei, A sharp stability result for the relative isoperimetric inequality inside convex cones, *J. Geom. Anal.*, Online First: 20 October 2011, to appear.
- [7] A. Figalli, F. Maggi, and A. Pratelli, A mass transportation approach to quantitative isoperimetric inequalities, *Invent. Math.*, 182 (2010), no. 1, 167-211.
- [8] A. Figalli, F. Maggi, and A. Pratelli, Sharp stability theorems for the anisotropic Sobolev and log-Sobolev inequalities on functions of bounded variation, *Adv. Math.*, to appear.
- [9] L. Gross, Logarithmic Sobolev Inequalities, *Amer. J. Math.*, 97 (1975), 1061-1083.
- [10] A.V. Kolesnikov, On Sobolev Regularity of Mass Transport and Transportation Inequalities, *arXiv:1007.1103v3*.
- [11] F. Otto and C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.*, 173 (2000), 361-400.
- [12] G. Perelman, The entropy formula for the Ricci flow and its geometric applications, *arXiv:math/0211159v1*.
- [13] G. Royer, An Initiation to Logarithmic Sobolev Inequalities. Translated from the 1999 French original by Donald Babbitt. *American Mathematical Society*, Texts and Monographs, 14, 2007.
- [14] A.J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Indo. Control* 2 (1959), 101-112.
- [15] C. Villani, Topics in optimal transportation. *Graduate Studies in Mathematics*, 58, American Mathematical Society, 2003.
- [16] C. Villani, Optimal transport, old and new. *Grundlehren der Mathematischen Wissenschaften* [Fundamental Principles of Mathematical Sciences], 338. Springer-Verlag, 2009.

EMANUEL INDREI

DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF TEXAS AT AUSTIN
1 UNIVERSITY STATION, C1200
AUSTIN TX 78712, USA
EMAIL: eindrei@math.utexas.edu

DIEGO MARCON

CENTER FOR MATHEMATICAL ANALYSIS,
GEOMETRY, AND DYNAMICAL SYSTEMS
DEPARTAMENTO DE MATEMÁTICA
INSTITUTO SUPERIOR TÉCNICO
LISBOA 1049-001, PORTUGAL
EMAIL: dmarcon@math.utexas.edu